EPA Feedback on LWG's 2/23/2011 MNR Presentation

General Comments:

The presentation was well organized and provided useful background on empirical lines of evidence that could be used to support MNR predictions. Model output was presented for contaminant levels in water and sediments for several contaminants over the length of the site. It is clear that the model is quite developed. EPA is providing the following observations and feedback of the government team reviewers:

.

1. There was little information on how the model's analyses actually generated results, in particular the derivation of future sediment contaminant concentrations. It would be quite useful to obtain an explanation in general terms and mathematically how sediment concentrations are determined over time for specific areas. In particular, what are the contaminant concentration terms used to establish t=0 concentrations, what concentration terms are used to derive sediment concentrations into the future, and what processes are modeled to feed those concentration terms.

2. Ultimately, potential remedial approaches will be evaluated on AOPC or SMA basis. It would be useful if an example AOPC was used to demonstrate how future sediment concentrations are depicted in the MNR model. The best example would be a PCB-contaminated AOPC that appears to be a good candidate for MNR. Such information needs to demonstrate: 1) for one representative model cell, how initial sediment concentrations are established along with the inputs, sources of those inputs, and equations used to predict future concentrations in that cell; 2) how information from individual cells within an AOPC will be combined to make statements about the magnitude and rate of MNR processes; and 3) how information from 1) and 2) will be compared to empirical data to ascertain whether model output appropriately represents MNR processes at the AOPC and hence is appropriate for decision making. Such information would permit a more rigorous evaluation of the MNR model. As of now, we have information on a few of the inputs and a depiction of output from the tool across the site, but we do not understand how results were generated within the model.

3. A major issue remains that it is uncertain whether the levels of precision and certainty that the model can achieve are sufficient to support decision making. It may well be that the model is the best one there is, but that doesn't mean it is a reliable tool for risk management decisions. As noted during the presentation, as a minimum, the uncertainty ranges need to be carried through the model outputs to clearly indicate the level of imprecision in the model. Further it was also noted that it might be useful to at least augment the model with a concerted field study of sedimentation and chemical changes over the time needed to design and implement a remedy.

4. It was not clear that the data presented provide as clear a picture of the river, especially in the AOPCs, as "highly depositional." The consistency and comparability of the data were not established, nor was the variability in the data accounted for, nor the applicability of general trends to specific locations. LWG should do a comparison of the erosional and depositional areas and the model predictions.

5. Sediment transport is episodic, which makes interpretation of historical data, from cores for example, and prediction of future conditions, such as site-specific deposition rates, very difficult.

6. The model needs to make a number of critical assumptions and ignore other potentially important factors that may affect sediment movement, e.g., bedload transport and propeller wash, any of which may make the modeling results not predictive of future trends. At some point, these issues need to be discussed.

7. The data available to calibrate the model are limited in some cases, e.g., the water column data, and highly variable in other cases, e.g., the surface sediment grain-size and chemical concentration data. These data make it difficult to constrain the model and demonstrate either it's accuracy or precision.

8. While numerous data were presented that indicated that sediment accretion was occurring in many areas, there was no attempt to more rigorously compare the trends indicated by the different types of data to determine whether they were consistent over time or space with each other and, more specifically, with the model outputs. For example, the comparison of surface and subsurface sediment concentrations were stated to indicate general decreases in concentrations over the last decade on the order of two- to five-fold. In comparison, the temporal trends in surface data are much larger.

9. The model is not an absolute predictor of sediment concentrations and time to cleanup. It is a tool that will be used to help compare various cleanup options. For instance, the MNR model may predict that it will take 30 years for the surface sediment at the site to meet background levels, however, if 30% of sediment is dredged (hottest material), then the MNR model may predict that it would take 15 years for surface sediment at the site to meet background. This would justify active remediation vs. MNR only. It also would show that we're looking at maybe 30-50 years to meet background vs. 500 years. The model won't be able to do much more than that. This model is too gross a scale to warrant fine precision, even to AOPC levels.

10. Caution must continue on the limits and boundaries of the model application and utility. The model alone should not be relied upon solely to make cleanup decisions and should be one line of evidence in how the system is behaving.

11. The model will only be useful insofar as it accurately reflects the interpretation of the data and its conclusions are accepted. This will likely still require a long term monitoring strategy for sediment, water and biota and will be necessary given the uncertainties that are involved. It will also be desirable given the improvements that are occurring in measurement technology; current measurement technology can record changes in sediment more reliably than modeling. A typical long-term monitoring program would likely involve an intensive sampling event that recurs every five year. The analysis of the data should include comparisons of the historical hydrograph and the actual bathymetry changes to see if the river bed is responding as expected. Supplemental sampling events associated with interim actions would be needed to look at active areas of remediation as the actions occur over the next 20-40 years.

12. Regarding the timeframe chosen as the basis for inputs to the analysis, it appeared from mapping that there was a considerable amount of deposition in the period used as the basis for the modeling. With low influent concentrations of contaminants in sediments entering the system to deposit, this would provide a favorable concentration profile for MNR.

The appearance of disproportionate levels of deposit compared to averages over a longer time frame may simply be the result of the 2-dimensionl representation of the information. Previously, information was presented on the results of a mass balance that looked at incoming suspended sediment load and exiting suspended sediment load that was used to support the contention that a certain amount of material was depositing in the study area. This is another line of evidence.

13. Source control: The information does not discuss the time frame for getting sources controlled and whether that timeframe is sufficient.

14. Grain size: the predominance of finer grained sediments is generally indicative of depositional environments. However, coarser materials are present in the RM 5 – 7 reach.

15. Surface vs. deep concentrations: Surface versus deep plots show that surface sediments are higher relative to subsurface sediments adjacent to nearby sources (PCBs). For DDx, the presence of NAPL at depth overwhelms the surface data; other locations are likely indicative of historic watershed-wide DDx sources. Caution should be used in making assumptions about subsurface vs surface concentrations when these conditions exist.

16. Temporal trend lines: The trend lines presented for PAHs are strongly influenced by early 1990's data. The data collected since the late 1990's seems relatively flat.

17. Depending on how locations for sediment traps were chosen, estimated sedimentation rates may be biased high relative to average site conditions.

18. The model looks at the "current-case" scenario for storm water inputs. As part of the sensitivity analysis on the MNR model, the LWG needs to look at source control reductions in storm water loading and include that in the FS report. EPA also requests that the LWGs consultants run the upper/lower boundary scenarios for stormwater and provide the agencies with the output when it has been completed (prior to submittal of the FS report).

Specific Comments:

Slide 7 names identification and characterization of ongoing sources as a data need for predicting MNR and states that LWG is using the RI report source table for this. This analysis also needs to consider future predicted sources, which are not provided in the RI report which could result in increases in storm water loading. Slides 70 & 73 say that other loading assumptions may be modeled. We agree this is necessary and should be part of a sensitivity analysis of the MNR model.

The information presented in the Table appears reasonable, however, the information presented over generalizes in some cases as highlighted in italics below. Specifically:

- Source control efforts currently underway at more than 80 upland sites – *Source control at these sites are at various stages in the source control process, but not complete. Since the degree of source control achieved and the timing of source control efforts is uncertain, this uncertainty needs to be accounted for in the analysis.*

- Several data sets evaluated consistently indicate the Study Area is largely depositional – *While it is true that the study area is generally depositional, it is not uniformly depositional. For example, the area between RM 5 – 7 is not necessarily depositional and is where some of the most significant sources are present.*

- Large areas of sediment bed are fine-grained, and likely relatively "stable" – *Many areas but not all are relatively stable. Again, the area between RM 5 and 7 is less stable than other areas of the site. The model may not accurately predict conditions in this area of the river.*

Slides 12-14: Bathymetric surveys often have an imprecision of 15 to 30 cm. It would be helpful to have the precision of the surveys more explicitly presented, particularly the longitudinal and cross-channel bed elevation plots. Similarly, the uncertainty in the estimated sedimentation rates should be stated. It might also be interesting to present the same figure and plots with the data converted to sedimentation rate instead of bed elevation change.

Slide 13 shows *laterally averaged* data on sedimentation rates. This is not very relevant since the nearshore sedimentation is significantly different from the channel sedimentation (as shown in Slides 12, 14, and 15) and the highest contamination is near shore. Therefore, in the FS, an area of the river should not be proposed for MNR based on cross-channel averages, but on laterally differentiated data.

Slide 18: From the brief description, the core data seem consistent with no or very low sedimentation.

Slide 20: Hill and McClaren (2001) also took a detailed survey of the surface sediment texture throughout the study area in 2000. It would be interesting to compare the data and it might also be helpful to use the 2000 data in the model.

Slides 21-24: Historical dredging areas and depths of dredging need to be included in evaluations of core lithology and historical events in the FS report, such as those presented in these slides.

Slide 22: It was hard to tell from the shot on the screen, but it appeared that there might be banding in Core C019, indicative of event-based (annual?) changes in deposition. Such data, if present, should be considered.

Slide 23: It might be noted that Core C002 was collected off shore of a sand and gravel operation. Spills and losses from overfilling of barges are not uncommon at such operations, and should be considered a possible source of the coarse material in that core.

Slides 23 and 24: It seems that at least some of the coarser material in cores C002 and C028 are large enough to be difficult to move as suspended matter. If so, this information should also be discussed in considering the uncertainties of the model.

Slide 27: It is mathematically disconcerting since ratios of (positive) concentrations can't be negative. It was initially thought the vertical scale was showing logarithms but for slide 28 logarithms do not seem correct because the DDx concentrations can be measured over 35 orders of magnitude. It is unclear as to whether these are actual ratios of concentrations or the ones that increase with depth have been flipped upside down. Further, the space between -1 and 1 on this chart is misleading since it can't be occupied, and the caption on the vertical scale is wrong for the "negative" values. These slides should be modified to accurately correspond to the mathematical definitions being used.

Slide 27-28: It is understood that the comparison of surface to sub-surface sediment was meant as a fairly general presentation, but these are the best data to evaluate whether the model is realistic. These data should be presented in more detail, e.g., the data from each core by river mile. In addition, the uncertainties in data should be considered, as well as providing a comparison of other data that might help interpret the chemical distributions, such as the percent fines in the different horizons, the distribution of other substances not expected to change over time, etc.

Slide 32: Without more information, the data presented in these plots can be misleading. The data collected on different surveys were often directed to substantially different locations to sample for different reasons. For example, the changes over time in concentrations indicated in these plots seem too much greater than those indicated by comparison of surface to sub-surface concentrations in the same locations. These data may be important it speaking to the accuracy of the model, but they need more explanation.

Slide 33: Note that the sediment samples used to evaluate temporal trends specifically exclude nearshore data because it could be confounded by nearshore sources. Since slides 12-15 show that sedimentation is significantly different in the nearshore versus the channel, any trends identified with this limited dataset cannot be extrapolated to the nearshore.

Slide 35: The data shown in the plots associated with McCormick and Baxter seem to show no changes over time if the earliest data are excluded. These data were collected by a different program and may have an unrecognized bias or other source that makes them not comparable. At a minimum, the lack of more recent changes in concentration should be noted.

Slides 30-37: Sediment chemical temporal trends. The speaker and presentation were clear that the data on temporal declines were heterogeneous and “mask out” temporal changes, that sediment sampling locations over time did not lend themselves to temporal analysis, that very few data points existed from the earliest time points, and that examples showing trends were cherry picked from the larger analysis. This is understandable recognizing the nature of sediment data and collection efforts. What is unclear is the resulting conclusion in slide 37 that the analysis shows evidence of temporal declines. Based on the presented information and caveats, a more defensible conclusion is that the analysis does not demonstrate trends (much less declines) and is not capable of demonstrating temporal declines.

Slide 42: It would be helpful to see plots of the concentration of the selected substances measured in non-SWAC areas, perhaps by river mile, to compare those data to the “source” concentrations.

Slides 42-44: In general it appears that, of the data shown, only PCBs and B(a)P were significantly lower in concentration in the upstream data compared to the average concentrations in most of the AOPCs. The concentrations of DDE in the particulates appeared to be higher than

observed at most of the AOPCs. These data indicate that the selection of the boundary concentration is important and not well established.

Slide 44: It is unclear why we are comparing incoming sediment chemistry to bed chemistry *in AOPCs* if the AOPCs are the more heavily contaminated areas (i.e. places where we will likely require active remediation, NOT MNR. In those areas, even if the incoming concentration is low, the existing bed concentration is hot enough that we likely need to remove it soon, rather than wait for natural recovery). If some of the AOPCs have lower concentrations such that MNR would be considered, those should be identified, and this analysis should focus on them and on the "in between" areas that are not part of any AOPCs.

Slide 46: It is noted that the east side traps are typically higher than west side, indicating that sources other than upstream (e.g. lateral sources like storm drains & CSOs) are important to consider in the model.

Slide 46: It would be helpful to re-plot these concentration data for the two sides of the river on the same scale, and also to test what appears to be systematic differences in the concentrations between the data from the east and west sides. If there is a difference, its importance to the modeling should be discussed.

Slides 63 on: The field data shown on these plots is highly variable and don't do a lot to constrain or help calibrate the model. Will additional model runs or data analyses, perhaps modeling smaller areas, be proposed to better understand the quality of the modeling results?

Slide 74: As noted on the call, please check the correlation among stormwater discharge, river discharge, and local precipitation, to potentially improve the accuracy of the model.

Slide 115: On the topic of including an active construction period before running the MNR model- on Duwamish, EPA instructed LDWG to model MNR during active construction because of concerns that not doing so would artificially inflate cleanup footprints and costs for the FS. For example, there may be some areas of the river that are currently slightly above cleanup standard concentrations and would merit active cleanup based on current concentrations, but would not be done until 5-10 years down the road, because the hotter spots would be prioritized first, or because of their location in the river. After 10 years of MNR, these areas might no longer need active dredging, so to include them in the FS makes that particular remedial alternative look artificially expensive and therefore less likely to be selected. To be realistic, the model should actually start now, or at the time of the last data point used for bed chemistry, rather than waiting.